Approximate Factorization of Multivariate Polynomials and Absolute Irreducibility Testing

AUTHOR(S):

Sasaki, T.; Suzuki, M.; Kolar, M.; Sasaki, M.

数理解析研究所講究録
第 746 巻 1991 年 59-68

Approximate Factorization of Multivariate Polynomials and Absolute Irreducibility Testing

T. Sasaki, M. Suzuki, M. Kolář and M. Sasaki

RIKEN (Institute of Physical and Chemical Research)
Wako-shi, Saitama, 351-01, Japan

Abstract

A concept of approximate factorization of multivariate polynomials is introduced and an algorithm for approximate factorization is presented. The algorithm handles polynomials with complex coefficients represented approximately, hence it can be used to test the absolute irreducibility of multivariate polynomials. The algorithm works as follows: given a monic square-free polynomial $F(x, y, \ldots, z)$, it calculates the roots of $F(x, y_0, \ldots, z_0)$ numerically, where $y_0, \ldots, z_0$ are suitably chosen numbers, then it constructs power series $F_1, \ldots, F_n$ such that $F(x, y, \ldots, z) \equiv F_1(x, y, \ldots, z) \cdots F_n(x, y, \ldots, z) \pmod{S^{e+2}}$, where $n = \deg_x(F)$, $S = (y - y_0, \ldots, z - z_0)$, and $e = \max\{\deg_y(F), \ldots, \deg_z(F)\}$; finally it finds the approximate divisors of F as products of elements of $\{F_1, \ldots, F_n\}$.

1 Introduction

Consider, for example, the polynomial

$$F(x, y) = x^3 + 3.0yx^2 + (2.2y^2 - 1.7)x - 2.9y.$$

Although $F(x, y)$ is irreducible, we can decompose it as

$$F(x, y) = (x^2 + 1.3yx - 1.7)(x + 1.7y) - 0.01y^2x - 0.01y.$$

We call this kind of decomposition approximate factorization of accuracy $\sim 10^{-2}$.

Factorization of polynomials has been studied by many persons. Some of the older algorithms are described in [vdW37]. The more modern ones which are practical but have exponential time-complexity in the worst case were devised by [Ber67, Zas69, WR75, WR76], etc. (see [Kal82a] for rather complete bibliography), and [LLL82, Kal82b, vzGK83, Len84], etc., have presented algorithms of polynomial time-complexity. Furthermore, several authors have studied factorization over an algebraically closed field and absolute irreducibility testing [HS81, CG82, Kal85, YNT90]. However, all of these studies deal with the exact factorization, hence the algorithms proposed are not applicable to polynomials with coefficients of floating-point numbers, for example.

On the other hand, scientists and engineers do often desire to "factorize" multivariate polynomials with coefficients of approximate numbers. According to their view, both $(x^2 - 10000)$ and $(x^2 - 10000.5)$, for example, are almost the same and "factorized" as

$$(x^2 - 10000.5) \simeq (x^2 - 10000) = (x + 100)(x - 100).$$

Thus, we are naturally led to extend the concept of factorization to the approximate one. This kind of extension is desirable not only for factorization but also for many other algebraic operations. In fact, recently one of the authors (T.S.) and his collaborators developed algorithms for approximate GCD, univariate and multivariate, and applied them to solving numerically ill-conditioned problems [SN89, SS90, ONS]. They called such algorithms the approximate algebraic algorithms which include algorithms for approximate factorization, too.

Since most of the conventional factorization algorithms employ the exact numerical arithmetic (to our knowledge, only [Len84] and [Kal85] employ the approximate numerical arithmetic), we cannot apply them to approximate factorization, and we need a new idea. Our idea is to use the fact that monic square-free polynomial $F(x, y)$ has the roots $x = \varphi_i(y)$, $i = 1, \ldots, \deg_x(F)$, of the form $\varphi_i(y) = c_{i,0} + c_{i,1}y + c_{i,2}y^2 + \cdots$. We thought that this fact had not been used in the conventional factorization algorithms, but we found that [Kal85] utilized this fact in a rather restricted way. However, we realized that the above-mentioned fact can be utilized more widely and elegantly than as used in [Kal85].

After giving some preliminaries in **2**, we reformulate in **3** the generalized Hensel construction which we will use to calculate the roots of $F(x, y)$ in the form of formal power series in y. The principle of (approximate) factorization is proved in **4**, and a primitive version of the algorithm is presented. The primitive version is, although complete, quite inefficient unless the degree of the polynomial is low. So, in **5**, we present a simple method for making the algorithm practical.

2 Preliminaries

In this section, we define notations, explain preprocessing of polynomials, and introduce notions on approximate factorization by referring to [SN89, SS90].

The field of complex numbers is denoted by C. By $\mathbf{C}[x, y, \ldots, z]$ and $\mathbf{C}\{y, \ldots, z\}[x]$ we mean, respectively, the polynomial ring over C in variables $x, y, \ldots, z$, and the ring of polynomials in x with coefficients of formal power series in $y, \ldots, z$ over C. In this paper, x is treated as the main variable.

Let $F(x, y, \ldots, z)$ be an element of $\mathbf{C}[x, y, \ldots, z]$ and

$$F(x, y, \ldots, z) = f_n(y, \ldots, z)x^n + f_{n-1}(y, \ldots, z)x^{n-1} + \cdots + f_0(y, \ldots, z), \quad f_n \not\equiv 0 \tag{2.1}$$

The *degree* of F w.r.t. x is denoted by $\deg_x(F)$ or simply by $\deg(F)$: $\deg(F) = n$. Similarly, the degree of F w.r.t. y is denoted by $\deg_y(F)$. The *leading coefficient* of F w.r.t. x is f_n and denoted by $\mathrm{lc}_x(F)$ or simply by $\mathrm{lc}(F)$: $\mathrm{lc}(F) = f_n$. Let f be a polynomial/power series in a single variable. By $\lfloor f \rfloor_e$, we mean the sum of all the terms, of f, of degrees not less than e: if $f = a_0 + a_1 y + a_2 y^2 + \cdots$, then $\lfloor f \rfloor_e = a_e y^e + a_{e+1} y^{e+1} + \cdots$.

Polynomial F is called *monic* w.r.t. x if $\mathrm{lc}_x(F) = 1$. Any polynomial can be transformed into a monic polynomial by the following transformation

$$F(x, y, \ldots, z) \to F'(x, y, \ldots, z) = f_n^{n-1} F(x/f_n, y, \ldots, z). \tag{2.2}$$

The $F(x, y, \ldots, z)$ can be decomposed as

$$\begin{aligned} &F(x, y, \ldots, z) = F_m^m(x, y, \ldots, z)F_{m-1}^{m-1}(x, y, \ldots, z) \cdots F_1(x, y, \ldots, z), \\ &\text{each } F_i \text{ has no multiple factor}, \quad i = 1, \ldots, m, \\ &\gcd(F_i, F_j) = 1 \quad \text{for } i \neq j. \end{aligned} \tag{2.3}$$

This decomposition is called *square-free decomposition* and each F_i is called *square-free*. Let $b_y, \ldots, b_z$ be numbers in C. If we choose $b_y, \ldots, b_z$ arbitrarily then $F(x, b_y, \ldots, b_z)$ is mostly square-free.

Remark 2.1 With the above preprocessing, we may assume without loss of generality that $F(x, y, \ldots, z)$ is monic and both $F(x, y, \ldots, z)$ and $F(x, 0, \ldots, 0)$ are square-free. □

Definition 2.1 [*maximum magnitude coefficient*]. The maximum magnitude numerical coefficient of F is denoted by $\mathrm{mmc}(F)$. ($\mathrm{mmc}(F)$ is nothing but the infinity norm $|F|_\infty$.) □

Definition 2.2 [numbers of similar magnitude]. Let a and b be numbers in C, and $b \neq 0$. By $a = \mathrm{O}(b)$, we mean that $1/c \leq |a/b| \leq c$ where $c \geq 1$ is a positive number not much different from 1. (Usually, "O" denotes Landau's notation, and we are using it in a somewhat different meaning.) □

Definition 2.3 [*normalization* of polynomial]. Normalization of polynomial F is the scale transformation $F \to F' = \eta F$, $\eta \in \mathbf{C}$, so that $\mathrm{mmc}(F') = 1$. □

Now, we define approximate factorization.

Definition 2.4 [*approximate factorization of accuracy* ε]. Let ε be a small positive number, $0 < \varepsilon \ll 1$. Let $F(x, y, \ldots, z)$ be a normalized polynomial in $\mathbf{C}[x, y, \ldots, z]$. If

$$F = GH + \Delta F, \quad \mathrm{mmc}(\Delta F) = \mathrm{O}(\varepsilon), \tag{2.4}$$

where G and H are non-constant elements of $\mathbf{C}[x, y, \ldots, z]$, then G (and H) is called an approximate divisor of F of accuracy ε. Finding G and H satisfying (2.4) is the approximate factorization, of accuracy ε, of F. □

Remark 2.2 If F is not approximately factorizable with accuracy less than $\mathrm{O}(\varepsilon)$ then F is irreducible and, in fact, absolutely irreducible because we are handling coefficients in C. □

3 Hensel construction in C{y}[x]

The Hensel construction is the p-adic univariate polynomial expansion technique devised by Hensel in ~1900, and it is generalized by Wang and Rothschild [WR75] to the expansion of multivariate polynomials. In this section, we formulate the generalized version in $\mathbf{C}\{y\}[x]$, for simplicity.

We assume that $F(x,y)$ is a monic polynomial in $\mathbf{C}\{y\}[x]$ and $F(x,0)$ is square-free. Let $\deg_x(F) = n$ and the roots of $F(x,0)$ be $u_1,\ldots,u_n$:

$$F(x,0) = (x-u_1)\cdots(x-u_n), \quad u_i \in \mathbf{C}, \quad i = 1,\ldots,n. \tag{3.1}$$

By assumption, $u_i \neq u_j$ for $i \neq j$. We put

$$F_i^{(0)} = x - u_i, \quad i = 1,\ldots,n. \tag{3.2}$$

Lemma 3.1 *For integer l, $0 \leq l \leq n-1$, there exists a set of numbers $\{w_1^{(l)},\ldots,w_n^{(l)}\}$ satisfying*

$$w_1^{(l)}[F_1^{(0)}\cdots F_n^{(0)}/F_1^{(0)}] + \cdots + w_n^{(l)}[F_1^{(0)}\cdots F_n^{(0)}/F_n^{(0)}] = x^l. \tag{3.3}$$

Proof Each term of the l.h.s. of this equation is a polynomial of degree $n-1$. Hence, we can determine $w_i^{(l)}$, $i = 1,\ldots,n$, uniquely by evaluating the above equation at n distinct points. Evaluating (3.5) at $x = u_1,\ldots,x = u_n$, we find

$$w_i^{(l)} = u_i^l / \prod_{j=1,\neq i}^{n} (u_i - u_j), \quad i = 1,\ldots,n. \tag{3.4}$$

□

If $H(x) = h_{n-1}x^{n-1} + \cdots + h_0x^0$, then (3.4) gives

$$\sum_{l=0}^{n-1} w_i^{(l)} h_l = H(u_i) / \prod_{j=1,\neq j}^{n} (u_i - u_j). \tag{3.5}$$

Lemma 3.2 [generalized Hensel's lemma]. *Let $F(x,y)$ be defined as above. For any positive integer k, there exists a set of polynomials $\{F_1^{(k)},\ldots,F_n^{(k)}\}$ satisfying*

$$F(x,y) \equiv F_1^{(k)}(x,y)\cdots F_n^{(k)}(x,y) \pmod{y^{k+1}}, \tag{3.6}$$

$$F_i^{(k)}(x,y) = x - u_i + c_{i,1}y + \cdots + c_{i,k}y^k, \quad i = 1,\ldots,n, \tag{3.7}$$

where $c_{i,j}$, $j = 1,\ldots,k$, are numbers in $\mathbf{C}$.

Proof By mathematical induction: Eqs.(3.6) and (3.7) are valid for $k = 0$. Assuming the validity of Eqs.(3.6) and (3.7), we consider the case of $k+1$. Put

$$F_i^{(k+1)} = F_i^{(k)} + \Delta F_i^{(k+1)}, \quad \Delta F_i^{(k+1)} \equiv 0 \pmod{y^{k+1}}, \quad i = 1,\ldots,n.$$

We determine $\Delta F_i^{(k+1)}$, $i = 1,\ldots,n$, so that they satisfy

$$F = (F_1^{(k)} + \Delta F_1^{(k+1)})\cdots(F_n^{(k)} + \Delta F_n^{(k+1)}) \pmod{y^{k+2}}$$

We can rewrite this equation as

$$\begin{aligned} \Delta F^{(k+1)} &\equiv F - F_1^{(k)}\cdots F_n^{(k)} \pmod{y^{k+2}}. \\ &= \Delta F_1^{(k+1)}[F_1^{(0)}\cdots F_n^{(0)}/F_1^{(0)}] + \cdots + \Delta F_n^{(k+1)}[F_1^{(0)}\cdots F_n^{(0)}/F_n^{(0)}]. \end{aligned} \tag{3.8}$$

By the induction assumption, the l.h.s. of this equation can be expressed as

$$\Delta F^{(k+1)} \equiv y^{k+1}(\tilde{f}_{n-1}x^{n-1} + \cdots + \tilde{f}_0) \pmod{y^{k+2}}. \tag{3.9}$$

Hence, by Lemma 3.1, we can determine $\Delta F_i^{(k+1)}$ as

$$\Delta F_i^{(k+1)} = y^{k+1}\Big(\sum_{l=0}^{n-1} w_i^{(l)}\tilde{f}_l\Big), \quad i = 1,\ldots,n. \tag{3.10}$$

The $F_i^{(k+1)}$ constructed is of the form of (3.7). □

Lemma 3.3 *(proof omitted). Let $F(x,y)$ be defined as above. Let*

$$F(x,y) = G(x,y) \cdot H(x,y), \tag{3.11}$$

where G and H are monic and elements of $\mathbf{C}\{y\}[x]$ in general. Then, the Hensel construction of F by the above mentioned method is equivalent to the Hensel constructions of G and H. □

4 Algorithm of approximate factorization

We first consider power series by Hensel construction. For simplicity, we confine ourselves to be in $\mathbf{C}\{y\}[x]$. Let F', $G^{(k)}$ and $H^{(k)}$ be elements of $\mathbf{C}\{y\}[x]$. If

$$F'(x,y) \equiv G^{(k)}(x,y)H^{(k)}(x,y) \pmod{y^{k+1}}, \tag{4.1}$$

for any positive integer k, we represent the case of $k = \infty$ as

$$F'(x,y) = G(x,y)H(x,y), \tag{4.2}$$

where $G = \lim_{k\to\infty} G^{(k)}$, $H = \lim_{k\to\infty} H^{(k)}$. We have seen in 3 that any polynomial $F(x,y)$ which is monic w.r t. x and for which $F(x,0)$ is square-free, is factorized in $\mathbf{C}\{y\}[x]$ as

$$\begin{aligned} F(x,y) &= F_1(x,y)\cdots F_n(x,y), \\ F_i(x,y) &= x - u_i + c_{i,1}y + \cdots + c_{i,k}y^k + \cdots, \quad i = 1,\ldots,n, \end{aligned} \tag{4.3}$$

where $u_i, c_{i,1}, \ldots, c_{i,k}, \ldots$ are numbers in $\mathbf{C}$. In this and following sections, we define

$$e = \deg_y(F). \tag{4.4}$$

We want to calculate every irreducible polynomial divisor G of F, where $G \in \mathbf{C}[x,y]$, in terms of $F_1,\ldots,F_n$. This is assured by the following theorem.

Theorem 4.1 *(proof omitted). Let F and F_i, $i = 1,\ldots,n$, be defined as above. Any polynomial divisor of F is a product of elements of $\{F_1,\ldots,F_n\}$, and each element F_i is contained in only one irreducible polynomial divisor of F.* □

Theorem 4.2 *Let ε be a small positive number, $0 < \varepsilon \ll 1$. Let F, G, H be normalized monic polynomials in $\mathbf{C}[x,y]$ such that $F(x,0)$ is square-free and*

$$\begin{aligned} F(x,y) &= G(x,y)H(x,y) + \Delta F, \\ F(x,0) &= G(x,0)H(x,0), \quad \mathrm{mmc}(\Delta F) = \mathrm{O}(\varepsilon). \end{aligned} \tag{4.5}$$

Let $F_1,\ldots,F_n$ be defined as in (4.3). *Then, some elements of $\{F_1,\ldots,F_n\}$, let them be $F_1,\ldots,F_m$ with $m < n$, satisfy*

$$\begin{aligned} &G(x,y) = F_1(x,y)\cdots F_m(x,y) + \Delta G(x,y), \\ &\Delta G(x,0) = 0, \quad \mathrm{mmc}(\Delta G) = \mathrm{O}(\varepsilon). \end{aligned} \tag{4.6}$$

Proof Put $F = \tilde{G}\tilde{H}$, $\tilde{G} = G - \Delta G$ and $\tilde{H} = H - \Delta H$, where

$$\tilde{G} = F_1\cdots F_m, \quad \tilde{H} = F_{m+1}\cdots F_n.$$

Substituting these into (4.5), we have

$$\Delta G \cdot H + \Delta H \cdot G + \Delta F = \Delta G \cdot \Delta H. \tag{4.7}$$

Formula (3.10) shows that the correction terms $\Delta F_i^{(k+1)}$, $i = 1,\ldots,n$, are linearly proportional to the coefficients of $\Delta F^{(k+1)}$. This means that $\Delta F_i^{(k+1)}$, $i = 1,\ldots,n$, go to zero uniformly as ΔF goes to zero, and so are ΔG and ΔH. Hence, (4.7) means that

$$\mathrm{O}(\mathrm{mmc}(\Delta G)) = \mathrm{O}(\mathrm{mmc}(\Delta H)) = \mathrm{O}(\mathrm{mmc}(\Delta F)) = \varepsilon.$$

□

If $G = F_1\cdots F_m$ is a polynomial divisor of F, then $G \equiv F_1^{(k)}\cdots F_m^{(k)} \pmod{y^{k+1}}$, where $F_i^{(k)}$, $i = 1,\ldots,n$, are the finite power series constructed by Hensel's method. In order to test whether G is an (approximate) divisor of F or not, we have only to set $k > e$ $(= \deg_y(F))$ and perform the trial division of F by G. Since $\deg(F_i) = 1, i = 1,\ldots,n$, we can find all the irreducible (approximate) factors of F by testing all the combinations of $F_1,\ldots,F_n$ successively from low to high degrees.

This observation leads us to the following primitive algorithm which is applicable to $F(x,y,\ldots,z) \in \mathbf{C}[x,y,\ldots,z]$.

Algorithm: Approx-Factorization (primitive version)

Input: a normalized monic and approximately square-free polynomial $F(x, y, \ldots, z)$,
ε = the accuracy of approximation, $0 < \varepsilon \ll 1$.

Output: $\{G_1, \ldots, G_r\}$, where each G_i is an irreducible approximate factor, of acccuracy less than ε, of F.

I. Choose constants $y_0, \ldots, z_0$ such that $F(x, y_0, \ldots, z_0)$ is approximately square-free, and calculate the roots of $F(x, y_0, \ldots, z_0)$ numerically:

$$F(x, y_0, \ldots, z_0) = (x - u_1) \cdots (x - u_n).$$

II. Put $S = (y - y_0, \ldots, z - z_0)$, an ideal generated by $y - y_0, \ldots, z - z_0$. Using the generalized Hensel algorithm, construct $F_1, \ldots, F_n$ such that

$$F(x, y, \ldots, z) \equiv F_1(x, y, \ldots, z) \cdots F_n(x, y, \ldots, z) \pmod{S^{e+2}},$$
where $e = \max\{\deg_y(F), \ldots, \deg_z(F)\}$. $i = 1, \ldots, n$.

III. Put $\Gamma := \phi$ (null set) and let $\tilde{\Gamma}$ be an ordered set of all the distinct factors of $F_1 \cdots F_n$, of degrees less than or equal to $n/2$, arranged in low-to-high degree order:

$$\tilde{\Gamma} := (F_1, \ldots, F_n, F_1F_2, \ldots, F_{n-1}F_n, F_1F_2F_3, \ldots),$$
$$\deg(\text{element of } \tilde{\Gamma}) \leq n/2.$$

While $\tilde{\Gamma} \neq \phi$ do
(G := [the first element of $\tilde{\Gamma}$];
if mmc(remainder(F, G)) $\leq O(\varepsilon)$
then $\Gamma := \Gamma \cup \{G\}$ and delete all the multiples of linear factors of G from $\tilde{\Gamma}$,
else $\tilde{\Gamma} := \tilde{\Gamma} - \{G\}$).

IV. Return Γ. □

5 Efficient selection of the relevant combinations

The factorization algorithm given in **4** is, although complete, quite inefficient unless $n = \deg(F)$ is small. The time-consuming step is step **III**: in the worst case, we must check 2^{n-1} different combinations of $F_1, \ldots, F_n$. In this section, we present a simple and efficient method for finding the relevant combinations of $F_1, \ldots, F_n$. Here, only a brief description of the method is presented, and the detailed analysis will be given elsewhere. Furthermore, we discuss in this paper only the properties of exact polynomial factors and the discussion from the viewpoint of approximate factorization will be given in another paper.

We assume again that F is in $\mathbf{C}[x, y]$ and, define $F_1, \ldots, F_n$ as in (4.3) and e as

$$e = \deg_y(F(x, y)). \tag{5.1}$$

The key point of our method is the investigation of the coefficients of higher degree terms in $F_1, \ldots, F_n$. The existence of such terms is assured by the following lemmas.

Lemma 5.1 *(proof omitted). Let $F(x, y) \in \mathbf{C}[x, y]$ and $F = GH$, where G and H are monic and elements of $\mathbf{C}\{y\}[x]$. If G is a finite power series in y, or $G \in \mathbf{C}[x, y]$, then G and H are polynomial divisors of F. (If G contains a term cy^k, $k > e$, then G is an infinite power series in y).* □

Our method of finding the relevant combinations is to utilize several kinds of linear relations satisfied by the coefficients of the power series factors of polynomials. The first relation is as follows.

Lemma 5.2 *Let F and $F_1, \ldots, F_n$ be defined as in* (4.3). *If*

$$G = F_1 \cdots F_m = x^m + g_{m-1}(y)x^{m-1} + \cdots + g_0(y)$$

is an element of $\mathbf{C}[x, y]$, *then*

$$c_{1,k} + \cdots + c_{m,k} = 0 \quad \textit{for any integer } k > e. \tag{5.2}$$

Proof Put $F_i(x, y) = x - \varphi_i(y)$, $i = 1, \ldots, n$, then $\varphi_1(y), \ldots, \varphi_m(y)$ are roots of $G(x, y) = 0$. Hence, from the relationship between roots and coefficients, we have

$$\sum_{i=1}^{m} \varphi_i(y) = -g_{m-1}(y). \tag{5.3}$$

Observing the y^k-term of this equation, we obtain (5.2). □

Terminology If $F_1, \ldots, F_m$ satisfy (5.2), we say $\{F_1, \ldots, F_m\}$ satisfies *zero-sum relation*. □

Let $M_1^{(e)}$ be the following matrix constructed from the coefficients of F_i, $i = 1, \ldots, n$:

$$M_1^{(e)} = \begin{pmatrix} c_{1,e+1} & c_{1,e+2} & c_{1,e+3} & \cdots \\ c_{2,e+1} & c_{2,e+2} & c_{2,e+3} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ c_{n,e+1} & c_{n,e+2} & c_{n,e+3} & \cdots \end{pmatrix} \tag{5.4}$$

Then, Lemma 5.2 leads to the following theorem.

Theorem 5.1 *The rank of matrix $M_1^{(e)}$ is equal to or less than $n - r$, where r is the number of irreducible polynomial divisors of F.* □

Corollary. If the rank of $M_1^{(e)}$ is $n - 1$ then F is irreducible over C, i.e., absolutely irreducible. □

If F contains irreducible factors of degree ≥ 2, then some rows of $M_1^{(e)}$ are nonnull by Lemma 5.1 and we can find nonzero elements of $M_1^{(e)}$ by constructing F_i^k, $i = 1, \ldots, n$, up to $k \geq 2e$. Therefore, in order to find the zero-sum relations, we need not calculate infinite power series but have only to calculate finite terms of F_i, $i = 1, \ldots, n$.

In many cases, we can find the relevant combinations of $\{F_1, \ldots, F_n\}$ by finding the zero-sum relations for row vectors of $M_1^{(e)}$. We first define non-overlapping relations.

Definition 5.1 [non-overlapping relations]. Let $v_1, \ldots, v_n$ be vectors in $\mathbf{C}^{n'}$, where $n \leq n'$, and $R : a_1 v_1 + \cdots + a_n v_n = 0$ and $R' : a'_1 v_1 + \cdots + a'_n v_n = 0$ be linear relations on $v_1, \ldots, v_n$, where $a_i \in \mathbf{C}$ and $a'_i \in \mathbf{C}$, $i = 1, \ldots, n$. R and R' are called *non-overlapping* if $|a_1 a'_1| + \cdots + |a_n a'_n| = 0$. □

If there are r linear relations which are linearly independent, then the relations form an r-dimensional vector space.

Since the polynomial divisors of F of the form $x - \varphi(y)$ can be found easily, we assume without loss of generality that every row of $M_1^{(e)}$ is nonnull. Let $M_1^{(e)} = (v_1, \ldots, v_n)^T$, then our problem is to find non-overlapping linear relations of the form $a_1 v_1 + \cdots + a_n v_n = 0$, where each a_i is either 1 or 0. The non-overlapping relations can be found by the following algorithm which finds all the linear relations first by neglecting the condition on a_i and then decomposes the relations to non-overlapping ones if possible. Note that there may not exist such non-overlapping relations.

Algorithm: Find-Relations

Input: Vectors $v_1, \ldots, v_n$ in $\mathbf{C}^{n'}$, $n \leq n'$;
Let **R** be the vector space of linear relations on $v_1, \ldots, v_n$, and let $r = \dim(\mathbf{R})$.

Output: If there exist r non-overlapping relations $\{\tilde{R}_1, \ldots, \tilde{R}_r\}$ then $\{\tilde{R}_1, \ldots, \tilde{R}_r\}$ else **NOT**.

I. Find all the linear relations on $v_1, \ldots, v_n$ by the Gaussian elimination.
Let the relations obtained be

$$R_i : a_{i1} v_1 + \cdots + a_{in} v_n = 0, \quad i = 1, \ldots, r.$$

II. Form the matrix A as follows:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rn} \end{pmatrix} \begin{matrix} \leftarrow R_1 \\ \vdots \\ \leftarrow R_r \end{matrix} \tag{5.5}$$

III. Eliminate the columns of A, from the 1st to $(r-1)$st, by the Gauss method (Gaussian elimination), then eliminate the columns backward, i.e., from the rth to 2nd.
Let the resulting matrix be $\tilde{A}$ which is of the form

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & & 0 & \tilde{a}_{1,r+1} & \cdots & \tilde{a}_{1n} \\ & \ddots & & \vdots & \ddots & \vdots \\ 0 & & \tilde{a}_{rr} & \tilde{a}_{r,r+1} & \cdots & \tilde{a}_{rn} \end{pmatrix} \begin{matrix} \leftarrow R'_1 \\ \vdots \\ \leftarrow R'_r \end{matrix} \tag{5.6}$$

IV. If $\tilde{A}$ consists of mutually non-overlapping row vectors then return the row vectors, else return NOT. □

Theorem 5.2 *Let $v_1, \ldots, v_n$ and r be defined as in algorithm Find-Relations. Then, the algorithm Find-Relations is correct, i.e., it finds all the non-overlapping linear relations on $v_1, \ldots, v_n$ provided that the number of such non-overlapping relations is r.*

Proof After the Gaussian elimination in Step I, all nonnull vectors calculated are linearly independent of each other, and $r = \dim(\mathbf{R})$ where $\mathbf{R}$ is the vector space of linear relations on $v_1, \ldots, v_n$. Next, we consider the matrix $\tilde{A}$ in (5.6). Note that $\tilde{a}_{11} \neq 0, \ldots, \tilde{a}_{rr} \neq 0$, because pivoting is made in Step I if necessary. We denote the ith row of $\tilde{A}$ by R'_i, $i = 1, \ldots, r$.

Suppose that there exist r non-overlapping linear relations $\tilde{R}_1, \ldots, \tilde{R}_r$ on $v_1, \ldots, v_n$. Then, we must have $\{R'_1, \ldots, R'_r\} = \{\tilde{R}_1, \ldots, \tilde{R}_r\}$, where the proportionality constants are omitted. In order to see this, suppose, for example, that $\tilde{R}_1 : \tilde{a}_{11}v_1 + \cdots + \tilde{a}_{1n}v_n = 0$ and $\tilde{R}_1 \neq R'_1$. Then, $\tilde{R}_1$ is expressed as $\tilde{R}_1 = R'_1 + b_2R'_2 + \cdots$, where at least one of b_i, $i = 2, \ldots, r$, is not zero. Suppose $b_2 \neq 0$, then $\tilde{R}_1$ is of the form $\tilde{R}_1 : \tilde{a}_{11}v_1 + b_2\tilde{a}_{22}v_2 + \cdots = 0$. Therefore, by assumption, $\tilde{R}_2, \ldots, \tilde{R}_n$ must not contain R'_1 and R'_2 so they must be expressed in terms of $R'_3, \ldots, R'_r$, which is impossible because $\tilde{R}_1, \ldots, \tilde{R}_n$ are linearly independent of each other. □

We show an example of factorization using the zero-sum relations for $M_1^{(e)}$.

Example 5.1 $F(x,y) = x^4 + (y-2)x^3 - (y+1)x^2 + (y^2+2)x - y$.
By Hensel construction, we obtain

$$\begin{aligned} F_1 &= (x-0) - y/2 - y^2/8 - y^3/16 - 5y^4/128 + \cdots, \\ F_2 &= (x-1) - y/2 - y^2/8 + 0 + y^4/128 + \cdots, \\ F_3 &= (x-2) + y/2 + y^2/8 + y^3/16 + 5y^4/128 + \cdots, \\ F_4 &= (x+1) - y/2 + y^2/8 + 0 - y^4/128 + \cdots. \end{aligned}$$

Coefficients of terms y^k, $k \geq 2$, give the matrix

$$M_1^{(2)} = \begin{pmatrix} -1/8 & -1/16 & -5/128 & \ldots \\ -1/8 & 0 & 1/128 & \ldots \\ 1/8 & 1/16 & 5/128 & \ldots \\ 1/8 & 0 & -1/128 & \ldots \\ -1/8 & -1/16 & -5/128 & \ldots \\ 0 & 1/16 & 6/128 & \ldots \\ 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & \ldots \end{pmatrix} \begin{matrix} \longleftarrow v_1 \\ \longleftarrow v_2 \\ \longleftarrow v_3 \\ \longleftarrow v_4 \\ \longleftarrow v_1 \\ \longleftarrow v_2 - v_1 \\ \longleftarrow v_3 + v_1 \\ \longleftarrow v_4 + v_2 \end{matrix}$$

We see that $\{F_1, F_3\}$ and $\{F_2, F_4\}$ satisfy the non-overlapping zero-sum relations, thus F_1F_3 and F_2F_4 are good candidates for polynomial divisors of F. In fact,

$$\begin{aligned} G_1 &\equiv F_1F_3 = x^2 - 2x - y, \quad G_1 | F, \\ G_2 &\equiv F_2F_4 = x^2 - yx - 1, \quad G_2 | F. \end{aligned}$$

□

Finally, let us consider the case in which, although we may determine some zero-sum relations, the above-mentioned column elimination operation on $M_1^{(e)}$ does not give us other zero-sum relations uniquely or some zero-sum relations are overlapping. In this case, we consider the zero-sum relations on the powers of roots. Let $\varphi_i(y)$ be the ith root of $F(x,y) = 0$:

$$F_i(x,y) = x - \varphi_i(y), \quad \varphi_i(y) = u_i - c_{i,1}y - c_{i,2}y^2 - \cdots, \quad i = 1, \ldots, n. \tag{5.7}$$

We calculate $\varphi_i^l, l = 2, \ldots, n$, and represent them as

$$\varphi_i^l = c_{i,0}^{(l)} + c_{i,1}^{(l)} y + \cdots + c_{i,k}^{(l)} y^k + \cdots, \quad l = 2, \ldots, n. \tag{5.8}$$

Without loss of generality, we assume that $G = F_1 \cdots F_m$ is a polynomial divisor of F. Then, $\varphi_1^l + \varphi_2^l + \cdots + \varphi_m^l$ can be represented by elementary symmetric polynomials in $\varphi_1, \ldots, \varphi_m$ and the degrees, w.r.t. y, of such symmetric polynomials are not greater than le. Hence, we have

$$\sum_{i=1}^{m} \lfloor \varphi_i^l \rfloor_{le+1} = 0, \quad l = 2, \ldots, n. \tag{5.9}$$

Thus, we again have zero-sum relations. Since the above relations must hold for each value of l independently, the above relations can be unified to the zero-sum relations on the combinations

$$t_1 \varphi_i + t_2 \varphi_i^2 + \cdots + t_n \varphi_i^n, \quad i = 1, \cdots, n, \tag{5.10}$$

where $t_1, t_2, \ldots, t_n$ are arbitrary parameters.

Let $\bar{M}^{(ne)}(t_1, \ldots, t_n)$ be the following matrix:

$$\bar{M}^{(ne)}(t_1, \ldots, t_n) = \begin{pmatrix} \sum_{l=1}^{n} t_l c_{1,ne+1}^{(l)} & \sum_{l=1}^{n} t_l c_{1,ne+2}^{(l)} & \cdots \\ \vdots & \vdots & \vdots \\ \sum_{l=1}^{n} t_l c_{n,ne+1}^{(l)} & \sum_{l=1}^{n} t_l c_{n,ne+2}^{(l)} & \cdots \end{pmatrix}. \tag{5.11}$$

Theorem 5.3 *If F has an irreducible divisor in $\mathbf{C}[x, y]$, let it be $G = F_{i_1} \cdots F_{i_m}$, then $\{F_{i_1}, \ldots, F_{i_m}\}$ satisfies the zero-sum relation for $\bar{M}^{(ne)}(t_1, \ldots, t_n)$. Conversely, if $\{F_{i_1}, \ldots, F_{i_m}\}$ satisfies the zero-sum relation for $\bar{M}^{(ne)}(t_1, \ldots, t_n)$, then $G = F_{i_1} \cdots F_{i_m}$ is a polynomial divisor of F.*

Proof The first half of the theorem is obvious from the above discussion, so we prove only the second half. By assumption, we see that for each l, $1 \le l \le n$, $\varphi_{i_1}^l + \varphi_{i_2}^l + \cdots + \varphi_{i_m}^l$ is a polynomial in y. By decomposing $\varphi_{i_1}^l + \varphi_{i_2}^l + \cdots + \varphi_{i_m}^l, l = 1, \ldots, n$, into elementary symmetric polynomials in $\varphi_{i_1}, \ldots, \varphi_{i_m}$, we see that the symmetric expressions are again polynomials in y. Hence, all the coefficients of $G = (x - \varphi_{i_1}(y)) \cdots (x - \varphi_{i_m}(y))$ are polynomials in y and G must be a polynomial divisor of F by Lemma 5.1. □

Example 5.2

$$\begin{aligned} F(x, y) = x^6 - & (12 - 12y + 3y^2)x^4 + (36 - 102y + 66y^2 - 30y^3 + 9y^4)x^2 \\ - & (81 - 198y + 247y^2 - 154y^3 + 49y^4). \end{aligned}$$

This example is constructed as $F = G_1 G_2$, where

$$G_1 = (x - 2 + y)^3 - (1 + y + y^2 + y^3), \quad G_2 = (x + 2 - y)^3 + (1 + y + y^2 + y^3).$$

Let $\omega_j, j = 1, \ldots, 6$, be the six different roots of $x^6 - 1 = 0$:

$$\omega_j = \cos((j - 1)\pi/3) + i \sin((j - 1)\pi/3), \quad j = 1, \ldots, 6.$$

Thus, we have $\omega_1 + \omega_4 = 0, \omega_2 + \omega_5 = 0, \omega_3 + \omega_6 = 0$. Let $\bar{\varphi}$ be

$$\begin{aligned} \bar{\varphi} &= (1 + y + y^2 + y^3)^{1/3} \\ &= 1 + y/3 + 2y^2/3^2 + 14y^3/3^4 - 46y^4/3^5 + 10y^5/3^6 + \cdots. \end{aligned}$$

Let the roots of $G_1(x, y)$ and $G_2(x, y)$ in $\mathbf{C}\{y\}[x]$ be $\{\varphi_1, \varphi_3, \varphi_5\}$ and $\{\varphi_2, \varphi_4, \varphi_6\}$, respectively:

$$\begin{aligned} \varphi_j &= 2 - y + \omega_j \bar{\varphi}, \quad j = 1, 3, 5, \\ \varphi_j &= -(2 - y) + \omega_j \bar{\varphi}, \quad j = 2, 4, 6. \end{aligned}$$

Thus, we have the relations

$$\varphi_1 + \varphi_4 = 0, \quad \varphi_2 + \varphi_5 = 0, \quad \varphi_3 + \varphi_6 = 0.$$

The F_i, $i = 1, \ldots, 6$, constructed from F by Hensel's method are

$$F_i(x, y) = x - \varphi_i(y), \quad i = 1, \ldots, 6.$$

Therefore, the rank of matrix $M_1^{(2)}$ is 1 and the row vectors of $M_1^{(2)}$ satisfy three zero-sum relations $v_1 + v_4 = 0$, $v_2 + v_5 = 0$, $v_3 + v_6 = 0$, as well as zero-sum relations $v_1 + v_3 + v_5 = 0$, $v_2 + v_4 + v_6 = 0$. However, only the latter two relations correspond to irreducible polynomial factors of F and the former three relations do not. Note that the above zero-sum relations are overlapping. (Applying the column elimination procedure mentioned above to $M_1^{(e)}$, we obtain $\omega_2 v_1 - v_2 = 0, \omega_3 v_1 - v_3 = 0, \ldots, \omega_6 v_1 - v_6 = 0$. These relations cannot be split into mutually non-overlapping linear relations by algorithm Find-Relations.)

On the other hand, we have

$$\begin{aligned}
\varphi_1^2 = \varphi_4^2 &= (2 - y)^2 + \omega_1(2 - y)\bar{\varphi} + \omega_1\bar{\varphi}^2, \\
\varphi_2^2 = \varphi_5^2 &= (2 - y)^2 - \omega_2(2 - y)\bar{\varphi} + \omega_3\bar{\varphi}^2, \\
\varphi_3^2 = \varphi_6^2 &= (2 - y)^2 + \omega_3(2 - y)\bar{\varphi} + \omega_5\bar{\varphi}^2.
\end{aligned}$$

Hence, eliminating the columns of matrix $\bar{M}^{(3)}(t_1, t_2, 0, \ldots, 0)$, we obtain only zero-sum relations

$$v_1 + v_3 + v_5 = 0, \quad v_2 + v_4 + v_6 = 0.$$

From these, we see that $G_1 = F_1F_3F_5$ and $F_2F_4F_6$ are good candidates for polynomial divisors of F, and in fact $G_1|F$ and $G_2|F$. □

6 Concluding remarks

In this paper, we have introduced the concept of approximate factorization, presented a primitive version of the algorithm of approximate factorization, and proposed a method which will make the algorithm practical. However, we have not fully analyzed the algorithm yet; a mathematical as well as computational analysis will be given elsewhere. In particular, we will show that finding the zero-sum relations on the matrix $\bar{M}^{(ne)}(t_1, \ldots, t_n)$ will determine the relevant combinations of $\{F_1, \ldots, F_n\}$ uniquely to give all the irreducible polynomial divisors of F [SSH91].

The basic principle of our factorization algorithm is quite simple, and we think the idea will be applicable to exact factorization over integers, etc. We are now pursuing such possibilities, too.

Acknowledgments

The authors thank Drs. T.Saito of Sophia Univ. and T.Hirano of Kanagawa Tech. Univ. for valuable discussions and Dr. K.Yokoyama of Fujitsu Ltd. for useful comments and references.

References

[Ber67] E.R. Berlekamp. Factorizng Polynomials over Finite Fields. *Bell System Tech. J.*, 46:1853–1859, 1967.

[CG82] A.L. Chistov and D.Y. Grigor'ev. Polynomial-time Factoring of Multivariate Polynomials over a Global Field. *LOMI preprint*, Lengingrad, 1982.

[HS81] J. Heintz and M. Sieveking. Absolute Primality of Polynomials is Decidable in Random Polynomial Time in the Number of Variables. In *Proc. 1981 Intn'l Conf. on Automata, Languages and Programming*, volume 115 of *Springer Lec. Notes Comp. Sci.*, pages 16–28. 1981.

[Kal82a] E. Kaltofen. Factorization of Polynomials. In B. Buchberger, G.E. Collins, and R. Loos, editors, *Computer Algebra: Symbolic and Algebraic Computation*, pages 95–113. Springer-Verlag, 1982.

[Kal82b] E. Kaltofen. A Polynomial Reduction from Multivariate to Bivariate Integral Polynomial Factorization. In *Proc. 1982 ACM Symp. on Theory Comp.*, pages 261–266. 1982.

[Kal85] E. Kaltofen. Fast Parallel Absolute Irreducibility Testing. *J. Symb. Comp.*, 1:57–67, 1985.

[Len84] A.K. Lenstra. Polynomial Factorization by Root Approximation. In *Proc. EUROSAM'84*, volume 174 of *Springer Lect. Notes Comp. Sci.*, pages 272–276. 1984.

[LLL82] A.K. Lenstra, Jr. H.W. Lenstra and L. Lovász. Factoring Polynomials with Rational Coefficients. *Math. Ann.*, 261:515–534, 1982.

[ONS] M. Ochi, M-T. Noda, and T. Sasaki. Approximate GCD of Multivariate Polynomials and Application to Ill-conditioned System of Algebraic Equations. (to appear in J. Inf. Process.).

[SN89] T. Sasaki and M-T. Noda. Approximate Square-free Decomposition and Root-finding of Ill-Conditioned Algebraic Equation. *J. Inf. Proces.*, 12(2):159–168, 1989.

[SS90] T. Sasaki and M. Sasaki. Analysis of Accuracy Decreasing in Polynomial Remainder Sequence with Floating-point Number Coefficients. *J. Inf. Proces.*, 12(4):394–403, 1990.

[SSH91] T. Sasaki, T. Saito, and T. Hirano, 1991. (preprint in preparation).

[vdW37] B.L. van der Waerden. *Moderne Algebra*, volume 1. Springer Verlarg, 2nd edition, 1937. translated to Japanese by K.Ginbayashi, Tokyo-Tosho, Tokyo, 1959.

[vzGK83] J. von zur Gathen and E. Kaltofen. A Polynomial-time Algorithm for Factoring Multivariate Polynomials over Finite Fields. In *Proc. 1983 Intn'l Conf. on Automata, Languages and Programming*, volume 154 of *Springer Lec. Notes Comp. Sci.*, pages 250–263. 1983.

[WR75] P.S. Wang and L.P. Rothschild. Factoring Multivariate Polynomials over the Integers. *Math. Comp.*, 29:935–950, 1975.

[WR76] P.J. Weinberger and L.P. Rothschild. Factoring Polynomials over Algebraic Number Fields. *ACM Trans. Math. Software*, 2:335–350, 1976.

[YNT90] K. Yokoyama, M. Noro, and T. Takeshima. On Factoring Multi-variate Polynomials over Algebraically Closed Fields, 1990. Proc. ISSAC'90 (to appear), ACM.

[Zas69] H. Zassenhaus. On Hensel Factorization I. *J. Number Theory*, 1:291–311, 1969.